z e t e r a

Zetera White Paper on

µSAN™ RAID

An overview, compare and contrast, technical paper on using the µSAN™ multicast protocol to implement different levels of RAID configuration.

This document references the µSAN™ White Paper (Version .35) and assumes the reader is familiar with the protocol.

Version 0.10

Author:
Charles W. (Bill) Frank
CTO, Zetera Corporation

January 2003

Revision History

Version 0.1	October 2002	Original Release

Contents

Overview 4

Overview

The μSAN™ RAID White Paper identifies a storage system and protocol wherein RAID configurations are established through the use of μSAN™ storage appliances connected by appropriately configured wired or wireless IP networks. By using optional RAID subset extension commands under multicast IP protocol, data can be presented to an array or fabric of μSAN™ storage appliances in stripes associated to sets of data. It is possible to establish RAID sets of types 0, 1, 4, 5 and 10 using the same topology since each μSAN™ participates autonomously within its set. This is an important advantage made possible by the combination of the autonomy of the μSAN™ and the ability of the multicast protocol to define groups of participants. Performance is scalable as a strong function of the bandwidth and capabilities of IP switching and routing elements and the number of participating μSAN™ appliances.

Basic definitions of RAID types:

RAID 0 — Data is striped across drives; no data redundancy is provided. Multiple disks are used to improve performance, but there is not logic to

RAID 1 — [illegible] disks at the same time. Data may be read from either disk, based on device availability. Although reliability is high, twice the amount of disk storage must be purchased.

RAID 2 — Data is striped at the bit level; multiple error-correcting disks provide redundancy; not a commercially implemented RAID level.

RAID 3 — Data is striped at the byte level, and one drive is set aside for parity information. This does well for large files where large blocks size and sequential I/O are used; not practical for small arrays.

RAID 4 — Data is striped in blocks, and one drive is set aside for parity information. This type is good for large block transfers; the parity drive is overworked if many small transfers are attempted-low IOPs.

RAID 5 — Data is striped in blocks, and parity information is rotated among all drives in the array. This type is best for small transfers-high IOPs.

RAID 10 — Data is striped across drives and mirrored on another striped set. The advantage is that the striped sets provide higher bandwidth access to large blocks of data. This definition is a popular construct not part of the original Berkley RAID definitions.

Multicast Data